JOC Articles

Single-Site Mutation and Secondary Structure Stability: An Isodesmic Reaction Approach. The Case of Unnatural Amino Acid Mutagenesis Ala→Lac

Andrzej Stanisław Cieplak* and Nur Başak Sürmeli

Department of Chemistry, Bilkent University, 06800 Bilkent, Ankara, Turkey

cieplak@fen.bilkent.edu.tr

Received December 18, 2003

A method is described to evaluate backbone interactions in proteins via computational unnatural amino acid mutagenesis. Several *N*-acetyl polyalanyl amides (AcA_nNH_2) were optimized in the representative helical (3_{10}-, 4_{13}-, and a "hybrid" κ-helix, $n = 7, 9, 10, 14$) and hairpin (two- and three-stranded antiparallel β-sheets with type I turns $\beta\alpha\alpha\epsilon$, $n = 6, 9, 10$) conformations, and extended conformers of *N*-acetyl polyalanyl methylamides ($n = 2, 3$) were used to derive multistranded β-sheet fragments. Subsequently, each residue of every model structure was substituted, one at a time, with L-lactic acid. The resulting mutant structures were again optimized, and group-transfer energies ΔE_{GT} were obtained as heats of the isodesmic reactions: $AcA_nNHR + AcOMe \rightarrow AcA_xLacA_yNHR + AcNHMe$ (R = H, CH_3). These group-transfer energies correlate with the degree of charge polarization of the substituted peptide linkages as measured by the difference Δe in H and O Mulliken populations in HN−C=O and with the H-bond distances in the "wild-type" structures. A good correlation obtains for the HF/3-21G and B3LYP/6-31G* group-transfer energies. The destabilization effects are interpreted in terms of loss of interstrand and intrastrand H-bonds, decrease in Lewis basicity of the C=O group, and O···O repulsion. On the basis of several comparisons of Ala → Lac ΔE_{GT}'s with heats of the NH → CH_2 substitutions, the latter contribution is estimated (B3LYP/6-31G*) to range between 1.5 and 2.4 kcal mol^{-1}, a figure close to the recent experimental $\Delta\Delta G°$ value of 2.6 kcal mol^{-1} (McComas, C. C.; Crowley, B. M.; Boger, D. L. *J. Am. Chem. Soc.* **2003**, *125*, 9314). The partitioning yields the following maximum values of the electronic association energy of H-bonds in the examined sample of model structures (B3LYP/6-31G* estimates): 3_{10}-helix $D_e = -1.7$ kcal mol^{-1}, α-helix $D_e = -3.8$ kcal mol^{-1}, β-sheet $D_e = -6.1$ kcal mol^{-1}. The premise of experimental evaluations of the backbone-backbone H-bonding that Ala → Lac substitution in proteins is isosteric (e.g., Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314) is often but not always corroborated. Examination of the integrity of H-bonding pattern and φ_i, ψ_i distribution identified several mutants with significant distortions of the "wild-type" structure resulting inter alia from the transitions between *i*, *i* + *3* and *i*, *i* + *4* H-bonding in helices, observed previously in the crystallographic studies of depsipeptides (Ohyama, T.; Oku, H.; Hiroki, A.; Maekawa, Y.; Yoshida, M.; Katakai, R. *Biopolymers* **2000**, *54*, 375; Karle, I. L.; Das, C.; Balaram, P. *Biopolymers* **2001**, *59*, 276). Thus, the isodesmic reaction approach provides a simple way to gauge how conformation of the polypeptide chain and dimensions of the H-bonding network affect the strength of backbone-backbone C=O···HN bonds. The results indicate that the stabilization provided by such interactions increases on going from 3_{10}-helix to α-helix to β-sheet.

1. Introduction

Site-directed mutagenesis has been successfully used to probe participation of individual residues in binding, catalytic action, folding, and stabilization of proteins.[1,2] However, the interpretation of the apparent energy of interaction $\Delta\Delta G_{app}$, i.e., the change in free energy of equilibrium or activation caused by a single amino acid substitution, is often difficult due to the uncertainties concerning solvation, mutant's structural integrity, and unfolded-state ensemble.[3] An example is unnatural amino acid mutagenesis aiming at evaluation of the role of backbone−backbone C=O···HN bonds in stabilization of the native structure,[4−8] still a controversial issue,[9,10] and

(1) Fersht, A. R.; Matoushek, A.; Serrano, L. *J. Mol. Biol.* **1992**, *224*, 771. Fersht, A. R.; Serrano, L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 75. Fersht, A. R. *Curr. Opin. Struct. Biol.* **1995**, *5*, 79.

(2) Chakrabartty, A.; Baldwin, R. L. *Adv. Protein Chem.* **1995**, *46*, 141. Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, *30*, 153.

(3) Fersht, A. *Structure and Mechanism in Protein Science*; W. H. Freeman & Co.: New York, 1999; Chapter 15.

10.1021/jo0358372 CCC: $27.50 © 2004 American Chemical Society
Published on Web 04/13/2004

in molecular recognition.[11−19] In this method, incorporation of an α-hydroxy acid into the polypeptide chain replaces the peptide linkage NH group with the ester linkage O.[20,21] Some results of such mutations appear consistent with the data suggesting that the strength of backbone−backbone H-bonds increases on going from turn or 3_{10}-helix to α-helix to *β*-sheet:[22] for instance, the substitutions in N−H-acceptor peptide bonds are destabilizing by 0.9 kcal mol-[1] in the α-helix 39−50 in T4 lysozyme, and by 1.5−2.5 kcal mol^{-1} in the antiparallel *β*-barrel of Staphylococcal nuclease.[4] Unfortunately, the experimental $\Delta\Delta G°$ determined upon deleting or perturbing one member of a hydrogen bond pair does not provide a direct measure of the strength of a hydrogen bond. Rather, the $\Delta\Delta G°$ reflects the difference between the amide interactions in the folded and unfolded states, and the ester interactions in the folded and unfolded states, all in water. For the enzyme−inhibitor complexes, the apparent free-energy change $\Delta\Delta G^{\text{amide}\rightarrow\text{ester}}$ associated with the NH → O substitution is proposed to be affected by loss of the backbone H-bond, the differential dehydration energy of inhibitors in the free form, and the electrostatic and van der Waals interaction between the two oxygen atoms in the H-bond depleted complex:[15,16] $\Delta\Delta G^{\text{amide}\rightarrow\text{ester}} = \Delta G^{\text{O}\cdots\text{O}} + (\Delta G^{\text{amide}}{}_{\text{solvation}} - \Delta G^{\text{ester}}{}_{\text{solvation}}) - \Delta\Delta G^{\text{H-bond}}$. In either case, a number of assumptions concerning the degree of solvent accessibility at the site of mutation, solvation energies, energies of other dipole−dipole interactions in the local protein environment, etc. are required to interpret the magnitude of the destabilization effect.[4a] So far, there is no conclusive evidence relating to such assumptions. For instance, the NH → O and NH → CH_2 substitutions were found to have the same effect on stability of the BPTI complex with trypsin,[13] hence neither $\Delta G^{\text{O}\cdots\text{O}}$ nor ($\Delta G^{\text{amide}}{}_{\text{solvation}} - \Delta G^{\text{ester}}{}_{\text{solvation}}$) contribution seemed here significant.[15,16] The same conclusion in regard to $\Delta G^{\text{O}\cdots\text{O}}$ was reached by comparing mutations at the N and C termini of α-helix 39−50 in T4 lysozyme.[4a] On the other hand, there is a considerable difference in the effects of the NH → O and NH → CH_2 substitutions on the complexes of thermolysin with phosphorus-containing peptide analogues,[12] and vancomycin with Ac$^{\text{D}}$A$^{\text{D}}$A,[19] and it was proposed that it is the O···O repulsion, not the H-bond loss, that is responsible for the larger share of the reduced binding affinity.[19] Furthermore, the free-energy perturbation calculations indicated that reduction in the solvation energy can be a major factor in the case of thermolysin complexes.[23] The need to develop a better understanding of the effects of the Ala → Lac substitution is underscored by a number of "anomalous" results: the destabilization effect can be greater for the mutation at the N−H-acceptor bond than at the N−H-donor bond (0.9 kcal mol^{-1} at N-terminus of α-helix 39−50 in T4 lysozyme, Leu-39, but only 0.7 kcal mol^{-1} at C-terminus of the same T4 lysozyme helix, Ile-50),[4a] or even greater than at the N−H-acceptor/donor bond (2.5 kcal mol^{-1} at Leu-14 of Staphylococcal nuclease, but only 1.7 kcal mol^{-1} at Ser-44 in the middle of α-helix 39−50 in T4 lysozyme),[4ab] while the NH → O substitution at the P2 position of eglin c unexpectedly leads to enhancement in both stability and binding to several serine proteases.[16a]

One way to circumvent many of the difficulties of the above method is to examine the effect of a single-site substitution by quantum mechanical methods. Ab initio MO or density functional theory studies will allow separation of the continuous dielectric and specific solvation effects, and remove the unfolded-state ensemble from the thermodynamic cycle by estimating the strength

(4) (a) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314. (b) Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 7151. (c) Shin, I.; Ting, A. Y.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 12667.

(5) Beligere, G. S.; Dawson, P. E. *J. Am. Chem. Soc.* **2000**, *122*, 120.

(6) Nakhle, B. M.; Silinski, P.; Fitzgerald, M. C. *J. Am. Chem. Soc.* **2000**, *122*, 8105. Wales, T. E.; Fitzgerald, M. C. *J. Am. Chem. Soc.* **2001**, *123*, 7709.

(7) Low, D. W.; Hill, M. G. *J. Am. Chem. Soc.* **2000**, *122*, 11039.

(8) Seebach, D.; Mahajan, Y. R.; Senthilkumar, R.; Rueping, M.; Jaun, B. *J. Chem. Soc., Chem. Commun.* **2002**, 1598.

(9) Creighton, T. *Curr. Opin. Chem. Biol.* **1991**, *1*, 5. Murphy, K. P.; Gill, S. J. *J. Mol. Biol.* **1991**, *222*, 699. Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660. Myers, J. K.; Pace, C. N. *Biophys. J.* **1996**, *71*, 2033.

(10) Dill, K. A. *Biochemistry* **1990**, *29*, 7133. Yang, A. S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 351.

(11) Bramson, N. H.; Thomas, N. E.; Kaiser, E. T. *J. Biol. Chem.* **1985**, *260*, 15452. Thomas, N. E.; Bramson, H. N.; Miller, W. T.; Kaiser, E. T. *Biochemistry* **1987**, *26*, 4461.

(12) Bartlett, P. A.; Marlowe, C. K. *Science* **1987**, *235*, 569. Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 297.

(13) Groeger, C.; Wenzel, H. R.; Tschesche, H. *Int. J. Peptide Protein Res.* **1994**, *44*, 166.

(14) Searle, M. S.; Sharman, G. J.; Groves, P.; Benhamu, B.; Beauregard, D. A.; Westwell, M. S.; Dancer, R. J.; Maguire, A. J.; Try, A. C.; Williams, D. H. *J. Chem. Soc., Perkin Trans. 1* **1996**, 2781.

(15) Lu, W.; Qasim, M. A.; Laskowski, M., Jr.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673.

(16) (a) Lu, W.; Randal, M.; Kossiakoff, A.; Kent, S. B. H. *Chem. Biol.* **1999**, *6*, 419. (b) Lu, W.-Y.; Starovasnik, M. A.; Dwyer, J. J.; Kossiakoff, A. A.; Kent, S. B. H.; Lu, W. *Biochemistry* **2000**, *39*, 3575.

(17) Baca, M.; Kent, S. B. H. *Tetrahedron* **2000**, *56*, 9503.

(18) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. *Biochemistry* **2001**, *40*, 7092.

(19) McComas, C. C.; Crowley, B. M.; Boger, D. L. *J. Am. Chem. Soc.* **2003**, *125*, 9314.

(20) For the effect of a single amide-to-ester replacement on ion channel function, see: England, P. M.; Zhang, Y.; Dougherty, D. A.; Lester, H. A. *Cell* **1999**, *96*, 89. Jude, A. R.; Providence, L. L.; Schmutzer, S. E.; Shobana, S.; Greathouse, D. V.; Andersen, O.; Koeppe, R. E., II. *Biochemistry* **2001**, *40*, 1460.

(21) For a general review of the applications of the unnatural amino acid mutagenesis, see: Dougherty, D. A. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645.

(22) The $^{3h}J_{\text{NiCj}}$ interactions across those bonds, reported to correlate with hydrogen bond distances, isotropic N_i-H chemical shifts, and $^{1}J_{\text{NiC'i}}$ couplings, tend to be greater for *β*-sheet H-bonds than for α-helix H-bonds, and have not been observed in 3_{10}-helices (Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601. Cornilescu, G.; Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949. Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275. Juranić, N.; Macura, S. *J. Am. Chem. Soc.* **2001**, *123*, 4099. Juranić, N.; Moncrieffe, M. C.; Likić, V. A.; Prendergrast, F. G.; Macura, S. *J. Am. Chem. Soc.* **2002**, *124*, 14221). This is in accord (b) with early conclusions of the surveys of H-bonding geometry in high-resolution crystal structures of proteins (Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97); (c) with trends in amide I and III band shifts in IR (cf. Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 12048); (d) and with the results of thermodynamic analyses of proteins (Wintrode, P. L.; Makhatadze, G. I.; Privalov, P. L. *Proteins: Struct., Funct., Genet.* **1994**, *18*, 2). (e) The opposite conclusion has been reached in the studies of D/H amide isotope effect which is most significant in α-helical proteins, weaker for α/*β*-, and negligible for all *β*-proteins: Khare, D.; Alexander, P.; Orban, J. *Biochemistry* **1999**, *38*, 3918. Shi, Z.; Krantz, B. A.; Kallenbach, N.; Sosnick, T. R. *Biochemistry* **2002**, *41*, 2120. It has been pointed out, however, that D/H fractionation at protein backbone amides reflects restrictions or enhancements of specific vibrational modes by the H-bond 3D-environment that is, in general, H-bonding geometry, and is largely independent of H-bonding strength: Bowers, P. M.; Klevit, R. E. *J. Am. Chem. Soc.* **2000**, *122*, 1030.

(23) Bash, P. A.; Kollman, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R. *Science* **1987**, *235*, 574. Mertz, K. M.; Kollman, P. K. *J. Am. Chem. Soc.* **1989**, *111*, 5649.

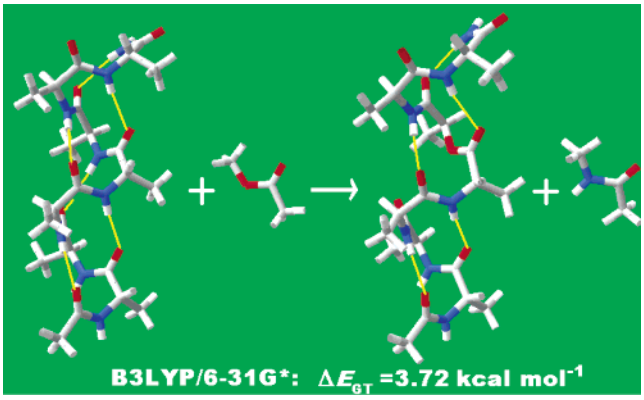
of a given bonding interaction via an isodesmic reaction scheme.[24] Furthermore, such studies might provide detailed information about the possible distortions of the wild-type structure that can be expected as a result of a single-site mutation.[25] In this paper, we describe an application of such a quantum chemical approach to the α-amino acid → α-hydroxy acid substitutions.

2. Computational Methods

In the first part of the study, a number of *N*-acetyl polyalanyl amides (AcA_nNH_2) were fully optimized in the conformations corresponding to the 3_{10}-(n=7, 9), 4_{13}-(α-helix, n = 10, 14), and "hybrid" κ-helix (an α-helix with the C terminal 3_{10}-turn, often encountered in proteins,[26] n = 9) as well as the hairpin and triple-stranded antiparallel β-sheets (type I turns $\beta\alpha\alpha\epsilon$, n = 6, 9, 10), at the HF/3-21G level of the theory. This method was recently reported by Topol et al. to yield satisfactory geometries of the local minima of *N*-formyl polyalanine amides such as 3_{10} and $4_{13}(\alpha)$ helices.[27] The protocol involved folding of the polyalanyl chain into the starting conformer using the standard φ_i and ψ_i values and subsequently an unconstrained optimization. As was reported earlier,[27] only in the case of some helical minima it was necessary to initially constrain the structure to preserve the desired H-bonding pattern through the early stage of optimization.[28,29] Thus, all the final helical and hairpin structures reported here were fully relaxed, and all the searches were completed by the default convergence criteria of Gaussian98. In the second part of the study, to derive planar antiparallel and parallel β-sheet models, the protocol of Kubelka and Keiderling was used,[22c] that is *N*-acetyl polyalanyl methylamides (AcA_nNHCH_3, n = 2, 3) were folded into extended strands using $\varphi_i = -138.6°$ and $\psi_i = 134.5°$ (the values from the crystal structure of β-sheet poly-L-alanine, for the planar antiparallel model) and $\varphi_i = -119°$ and $\psi_i = 113°$ (the standard values for the planar parallel model), and partially optimized with the φ_i and ψ_i torsional angles constrained to the above values. The strands were assembled into binary and tertiary complexes which were partially optimized, that is the φ_i and ψ_i angles were kept frozen at the initial level.

Subsequently, each residue of every "native" model structure was substituted, one at the time, with L-lactic acid, i.e., the backbone NH group was replaced by the O group. The resulting mutant structures were again optimized: full optimizations were performed in the case of the helical and hairpin conformers, and partial optimizations in the case of the multistranded β-sheet models; that is, the φ_i and ψ_i torsion angles continued to be constrained to the initial values.

SCHEME 1. Concept of Isodesmic Group-Transfer Reaction Applied to the Ala → Lac Mutagenesis[a]



[a] The scheme presents a hypothetical reaction of AcA_7NH_2 in 3_{10}-helix conformation **I** with methyl acetate which yields 3_{10}-helical depsipeptide ***I5*** and *N*-methyl acetamide. The mutation site is m = 5 (the residue and peptide bond numbering begins at the N-terminus at the bottom of the diagram). Since the type and number of covalent bonds in the educts and the products are the same, and the chain conformation is roughly preserved in the depsipeptide, the major effect on the potential energy is expected to be due to loss of one backbone-backbone H-bond (O2···NH5) and weakening of another one (O4···NH7), and O2···O5L repulsion.

The group-transfer energies ΔE_{GT} for the above substitutions were obtained as heats of the isodesmic reactions: AcA_n-NHR + AcOMe → AcA_xLacA_yNHR + AcNHMe (R = H, CH_3). The concept of such a reaction as it applies here is illustrated in Scheme 1.

To establish how reliable are the HF/3-21G group-transfer energies, a number of wild-type structures and their depsipeptide mutants were reoptimized at the B3LYP/6-31G* level (in addition, a few helical conformers were reoptimized at the HF/6-31G** and B3LYP/D95** levels). In each case, the optimization was continued untill the default convergence criteria were fully met. All of the calculations were performed using parallel version of Gaussian98 Revision A.7 installed on Sun Enterprise 4500 High-Performance Server, and Gaussian98 Revision A.11.2.[30]

3. Results and Discussion

a. Ala → Lac Substitutions in Helix Conformers. The group-transfer energies ΔE_{GT} for mutations of the fully optimized helical conformers are summarized in Table 1, and the representative native structures are shown in Chart 1. These reactions are endothermic

(24) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796. See also: Wiberg, K. B. *Acc. Chem. Res.* **1999**, *32*, 922.

(25) For the structural studies (X-ray) of depsipeptides, see: (a) Valle, G.; Bardi, R.; Piazzesi, A. M.; Crisma, M.; Toniolo, C.; Cavicchioni, G.; Uma, K.; Balaram, P. *Biopolymers* **1991**, *31*, 1669. (b) Crisma, M.; Valle, G.; Bonora, G. M.; Toniolo, C.; Cavicchioni, G. *Int. J. Peptide Protein Res.* **1993**, *41*, 553. (c) Ohyama, T.; Oku, H.; Hiroki, A.; Maekawa, Y.; Yoshida, M.; Katakai, R. *Biopolymers* **2000**, *54*, 375. (d) Ohyama, T.; Oku, H.; Hiroki, A.; Yoshida, M.; Katakai, R. *Biopolymers* **2001**, *58*, 636. (e) Karle, I. L.; Das, C.; Balaram, P. *Biopolymers* **2001**, *59*, 276. (f) Arawinda, S.; Shamala, N.; Das, C.; Balaram, P. *Biopolymers* **2002**, *64*, 255. (g) Peggion, C.; Barazza, A.; Formaggio, F.; Crisma, M.; Toniolo, C.; Villa, M.; Tomasini, C.; Mayrhofer, H.; Pöchlauer, P.; Kaptein, B.; Broxterman, Q. B. *J. Chem. Soc., Perkin Trans. 2* **2002**, 644.

(26) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601.

(27) Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T.-H.; Perczel, A.; Rashin, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **2001**, *123*, 6054.

(28) For the most recent full optimizations of the secondary structure models, see: (a) Perczel, A.; Jàkli, I.; Csizmadia, I. G. *Chem. Eur. J.* **2003**, *9*, 5332. (b) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, 8124. (c) Bour, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45.

(29) For a DFT study of polyglycine helix models based on the repeating unit approach, see: Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313. (b) For the studies of secondary structure models employing periodic boundary conditions, see: Improta, R.; Barone, V.; Kudin, K. N.; Scuseria, G. *J. Am. Chem. Soc.* **2001**, *123*, 3311. Rossmeisl, J.; Hinneman, B.; Jacobsen, K. W.; Nørskov, J. *J. Chem. Phys.* **2003**, *118*, 9783.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Rega, N.; Salvador, S.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; P. Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 2001.

TABLE 1. Group-Transfer Energies ΔE_{GT} (kcal mol^{-1}) for *m* Ala → Lac Mutations of *N*-acetyl Polyalanine Amides from the Isodesmic Reactions AcA_nNH_2 + AcOMe → $AcA_xLacA_yNH_2$ + AcNHMe (*m* Denotes the Mutation Site)[a]

m	3_{10}-helix		α-helix		κ-helix	hairpin	triple-stranded hairpin β-sheet	
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**
1	1.7	1.9	3.0	3.3	2.9	8.8	9.1	8.9
2	1.6	1.8	1.8	2.2	1.8	13.5	12.6	12.4
3	3.7*	2.6*	3.1	3.5	3.0	−0.9	−2.0	−0.6
4	7.6	8.3	5.0	5.9	5.0	0.3	0.2	0.6
5	7.3	8.2	5.7*	7.1*	5.5*	4.8*	2.6*	6.1*
6	7.2	8.4	8.2	9.9	7.7	9.0	13.2	14.8
7	4.0	8.1	9.8	11.4	8.7		6.5*	2.7*
8		7.6	9.4	11.6	9.9		0.6	0.4
9		4.6	5.7	11.7	9.7		6.9	7.1
10			5.2	11.2	6.4			9.1
11				11.4				
12				10.4				
13				6.5				
14				6.4				

[a] Asterisks indicate that the H-bonding pattern of the "wild-type" structure is not preserved in the depsipeptide mutant and/or a compensating donor−acceptor interaction is introduced.

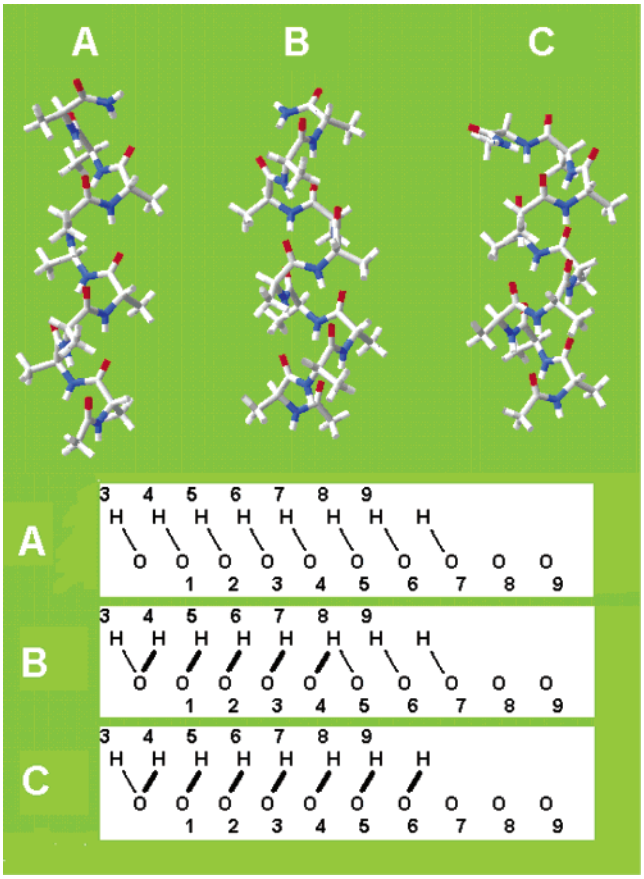
($\Delta E_{GT} > 0$) that is the effect of mutation is, as expected, destabilizing. In the case of the *endo* peptide bonds in AcA_3NHCH_3 constrained into a single α-helix turn, **IX**, the substitutions are thermoneutral; i.e., the NH and O interactions with the immediate molecular environment appear equivalent both in N,O-methyl acetyl derivatives and in the helical conformers of the peptide chain.[31]

Examination of the integrity of H-bonding patterns and ψ_i, φ_i distribution in the depsipeptide mutants reveals a wide range of distortions of the "wild-type" structure. On one hand, the Ala → Lac substitution often causes just a small "localized" change of the backbone torsional angles so as to be nearly isosteric; see ***I5*** in Scheme 1. On the other hand, this substitution can also produce a major conformational change associated with a formation of an additional, compensatory backbone−backbone interaction.

In the case of the helical structures, such distortions involve transitions from *i*, *i* + 3 to *i*, *i* + 4 H-bonding in the first turn of the 3_{10}-helix as a result of mutation at the site *m* = 3 (entries ***I3*** and ***II3*** in Table 1) and transitions from *i*, *i* + 4 to *i*, *i* + 3 H-bonding in the first turn of the α-helix as a result of mutation at the site *m* = 5 (entries ***III5***, ***IV5***, and ***V5*** in Table 1). The latter transition is reminiscent of the appearance of a 3_{10}-helical segment at the connective part between the peptide and the depsipeptide units in the crystal structure of a pentadecadepsipeptide $Boc(L_2A)_2(L_2Lac)_3OEt$ reported by Katakai et al.[25c] and at the mutation sites in depsipeptides BocVALAibVLacLAibVALOMe and BocVALAibVLacLAibVLOMe reported by Karle, Das, and Balaram.[25e] These transitions can be conveniently depicted in the H-bonding schemes in Chart 1: (i) in the diagram A, the N-terminal acetyl O (no label) would bond to N4H after the removal of N3H (mutation *m* = 3, followed by the transition from *i*, *i* + 3 to *i*, *i* + 4 H-bonding), (ii) in the diagrams B and C, the O1 atom would bond to N4H after the removal of N5H (mutation *m* = 5, followed by the transition from *i*, *i* + 4 to *i*, *i* + 3 H-bonding).

(31) **IX** is AcA_3NHCH_3 folded into a single turn of the α-helix using $\varphi_i = -63.4°$ and $\psi_i = -37.6°$, and partially optimized with all the φ_i and ψ_i torsional angles constrained to the above values. The ΔE_{GT}'s for **IX2**, **IX3** are −0.19, 0.63 kcal mol^{-1} (HF/3-21G), and −1.08, −0.29 kcal mol^{-1} (B3LYP/6-31G*).

CHART 1. Three Types of Right-Handed Helices Examined in the Present Study: (A) 3_{10}-Helix AcA_9NH_2 II; (B) κ-Helix AcA_9NH_2 V; (C) α-Helix $AcA_{10}NH_2$ III[a]



[a] The pattern of the backbone−backbone H-bonding in each helix is shown in the convention of Topol et al.[27] The residues (mutation sites) are numbered beginning at the N-terminus at the bottom of the diagram.

The extent of the "localized" conformational distortions is illustrated by the data in Table 2 which lists the backbone torsion angles in 3_{10}-helix AcA_9NH_2[28c] and its *m* = 6 mutant, $AcA_5LacA_3NH_2$, at different levels of theory. In accord with the previous report,[27,28a] there is a good agreement between the HF/3-21G and B3LYP/6-31G* geometries, with a minor change in the conformations of the terminal residues.

Examination of the molecular geometry of the mutant helices suggests that the main reason for the distortion is the O···O repulsion rather than the difference in the amide and ester torsional potentials. The O···N separation in helical backbone-backbone H-bonds is uniformly very close to 3.0 Å, but the O···O separation in depsipeptide 3_{10}-helices is 3.5−3.7 Å, and in 4_{13}-helices 3.2−3.3 Å (HF/3-21G; B3LYP/6-31G* distances are 0.1−0.2 Å greater). The corresponding distances in the crystal structures are reported to be 3.81(2), 3.87(2) Å,[25c] and 3.24(2), 3.47(2),[25d] 3.1−3.3 Å,[25e] respectively. Thus, the

TABLE 2. Main-Chain Torsional Angles of AcA_9NH_2 (II) and Its Mutant $m = 6$ (II*6*) in the 3_{10}-Helix Conformations at Different Levels of the Theory

residue no. i	HF 3-21G ψ_i	HF 3-21G φ_i	HF 6-31G** ψ_i	HF 6-31G** φ_i	B3LYP 6-31G* ψ_i	B3LYP 6-31G* φ_i	B3LYP D95** ψ_i	B3LYP D95** φ_i
AcA_9NH_2								
1	−30.2	−60.6	−25.6	−67.2	−25.1	−65.7	−25.3	−64.5
2	−22.7	−58.7	−21.7	−62.2	−21.6	−59.3	−21.9	−58.4
3	−21.8	−60.2	−22.0	−63.1	−20.8	−60.9	−21.5	−60.1
4	−22.1	−59.8	−21.8	−63.2	−22.3	−60.0	−22.0	−59.5
5	−21.5	−60.3	−21.5	−63.4	−20.7	−61.5	−20.7	−60.7
6	−20.6	−61.4	−20.8	−64.0	−20.3	−61.8	−20.6	−60.9
7	−21.2	−62.5	−20.6	−64.7	−20.6	−62.6	−20.4	−61.8
8	−3.0	−72.9	−12.9	−69.8	−7.7	−70.2	−10.8	−68.1
9	10.5	−103.5	2.2	−94.1	11.3	−102.6	3.3	−93.8
$AcA_3LacA_3NH_2$ m=6								
1	−29.7	−61.3			−23.9	−65.7		
2	−26.1	−58.5			−23.5	−59.5		
3	−6.1	−70.2			−12.1	−67.5		
4	−18.2	−76.9			−12.4	−80.4		
5	−33.3	−52.0			−25.2	−57.1		
6	−18.4	−65.0			−18.5	−66.9		
7	−21.3	−63.6			−20.1	−64.3		
8	−3.1	−72.7			−7.8	−70.0		
9	10.5	−103.6			10.9	−102.7		

depsipeptide helices are somewhat open on one side and compressed on the other, in the manner of selectively solvated α-helices in proteins.[26]

The effect of Ala → Lac substitution on charge distribution in the helices is illustrated in Figure 1 using the difference Δe in H and O Mulliken populations in HN−C═O as a measure of charge polarization of the peptide bond. The mutation at the site m causes, as expected, a decrease in charge polarization of the peptide bonds $m - 3$ and $m + 3$ (negative deviations in the plots in Figure 1), but it consistently increases the charge polarization of the immediately preceding $m - 1$ bond. The increase is apparently due to the bonding interaction between the peptide and ester linkages, revealed by the short C_i═O···C_j═O contacts (i = amide, j = ester; 2.7−2.8 Å rather than the standard 2.9−3.0 Å, HF/3-21G).[32,33] The NBO E(2) energies of the n_O(NC═O)−π*(OC═O) interactions in the 3_{10}-helix are indeed ∼2.5−3.0 times greater than those of the n_O(NC═O)−π*(NC═O) interactions (e.g., LP(2) O37−BD*(2) C45−O47 in **I**: 0.90 kcal mol^{-1}; in **I*5***: 2.62 kcal mol^{-1}, HF/3-21G).[34]

Smaller increases in charge polarization of the peptide bonds immediately following the mutation site can probably be attributed to the inductive effect which lowers Lewis basicity of the C═O group and increases Lewis acidity of the N−H group with the net change which usually is negligible, except for $m = 3$, 10, 12. The data shown in Figure 1 are obtained at the HF/3-21G level, but the changes in charge polarization of the peptide bonds appear to be quite well reproduced at this level of theory. The Δe values obtained for the reoptimized structures **I**, **I*4***, **II**, and **II*6*** are plotted against the corresponding HF/3-21G values in Figure 2. The anisotropy of charge distribution is exaggerated at the lower levels of theory; nonetheless, the HF/3-21G model is certainly qualitatively useful.[27,28a]

The data for the mutants with altered backbone−backbone interaction patterns are marked in Table 1 with asterisks. For all the other structures, heat of the substitution reaction is expected to reflect primarily loss of C═O···H−N bonding. It is therefore quite interesting to see in Table 1 a wide range of the ΔE_{GT} values for the helical polyalanines **I**−**V**: (i) 1.7−3.3 kcal mol^{-1} in the case of substitutions in the bonds that are exclusively N−H-acceptors, (ii) 4.0−10.4 kcal mol^{-1} if the peptide bonds are exclusively N−H-donors, and (iii) 5.0−11.7 kcal mol^{-1} for the peptide bonds in the middle of the trimeric or longer arrays, i.e., the bonds which are simultaneously N−H-acceptors and N−H-donors. There are two noteworthy trends in these data: the increase in ΔE_{GT} with increasing length of the given type helices, i.e., **I** vs **II**, or **III** vs **IV**, and the increase in ΔE_{GT} on going from 3_{10}-helix to 4_{13}-helix, e.g., **II** vs **III**.

b. Multiple Ala → Lac Substitutions in the α-Helix. The early studies of the sequential polydepsipeptides poly(L_2Lac) by Goodman et al. have suggested that while the substitution certainly decreases peptide helicity, even a multiple substitution does not entirely prevent folding into helical structures in nonpolar solvents.[35] Indeed, several Lac-based depsipeptides were recently found to adopt helical conformations by crystal structure analysis.[25c−f] Furthermore, chymotrypsin inhibitor 2 (CI2) has been shown to tolerate, in terms of the folding characteristics, a replacement of an array of four amide bonds that span the length of its α-helix with ester bonds.[5] To model the latter modification, four substitutions were introduced in *N*-acetyldodecyl amide **X**

(32) Bent, H. *Chem. Rev.* **1968**, *68*, 587.
(33) Bürgi, H. B.; Dunitz, J. D. *Acc. Chem. Res.* **1983**, *16*, 153. See also: Cieplak, A. S. In *Structure Correlation*; Dunitz, J. D., Bürgi, H.-B., Eds.; Verlag Chemie: Weinheim, 1994; Vol. 1, Chapter 6, pp 205−302.
(34) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735. Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO Version 3.1.
(35) Ingwall, R. T., Goodman, M. *Macromolecules*, **1974**, *7*, 598. Wouters, G.; Katakai, R.; Becktel, W. J.; Goodman, M. *Macromolecules* **1982**, *15*, 31, and references therein.

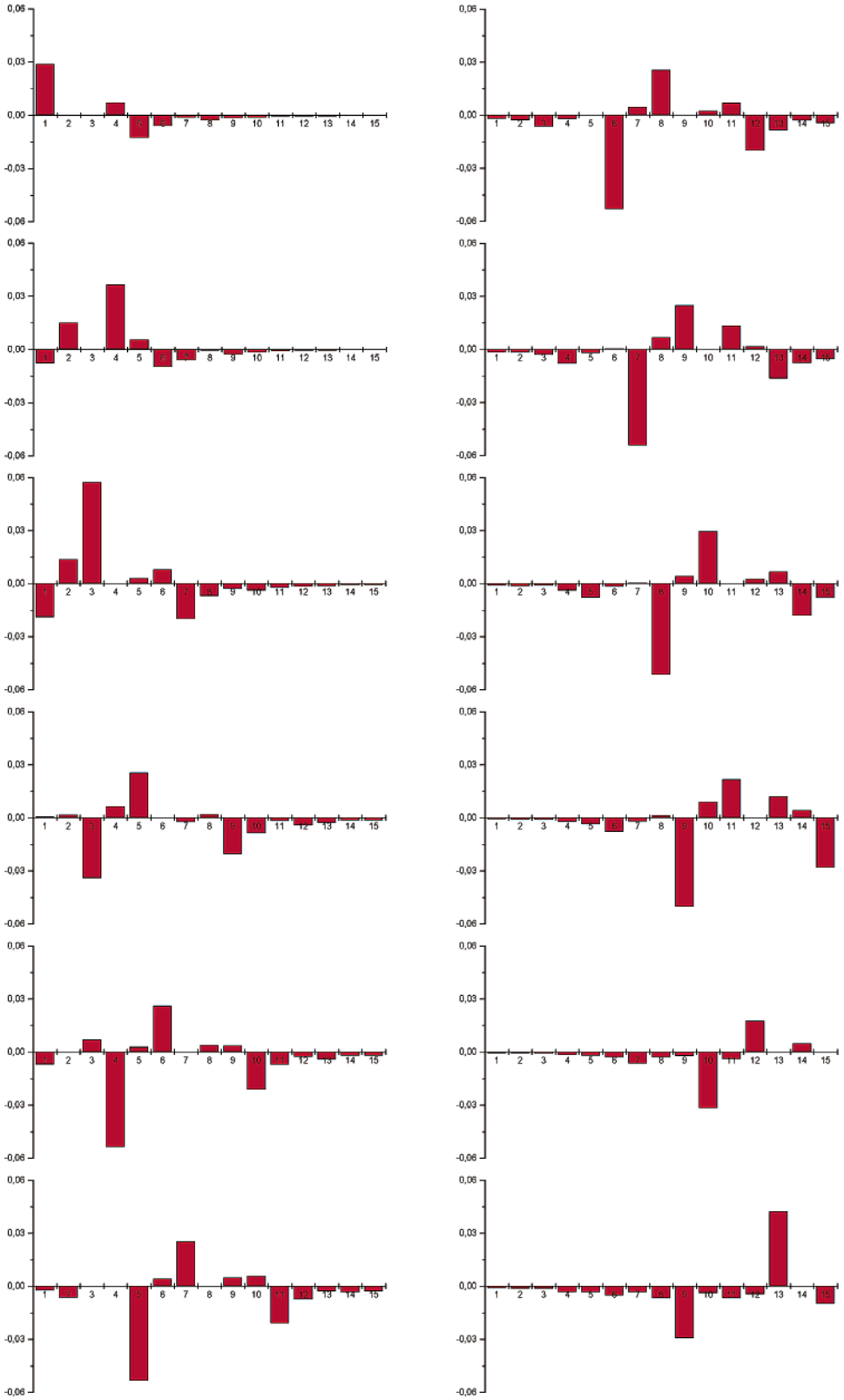


FIGURE 1. Difference $\Delta\Delta e$ in charge polarization Δe (see text) of the peptide bonds (HF/3-21G) in the depsipeptide mutants **IV*m*** and in the "wild-type" helix **IV**, $\Delta\Delta e_i = \Delta e_i(\mathbf{WT}m) - \Delta e_i(\mathbf{WT})$ (au), as a function of the bond location along the peptide chain i ($m = 2-4, 6-14$). A positive deviation indicates that the peptide bond is more charge-polarized in the mutant.

($AcA_{12}NH_2$) to yield $AcALacA_2LacA_2LacA_2LacANH_2$ ***X2.5.8.11***, and ΔE_{GT}=26.1 kcal mol^{-1}. This value exceeds somewhat the anticipated destabilization effect of ~18 kcal mol^{-1} (loss of three H-bonds), but the 4_{13}-helical conformation is preserved in the quadruple mutant, Chart 2. The C═O···O separations are about 3.1 Å. Interestingly, H-bonds in the remaining two amide arrays in the mutant are considerably shorter than the

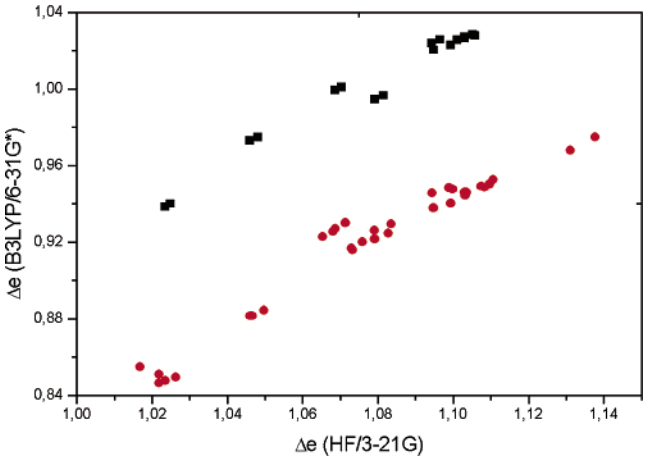
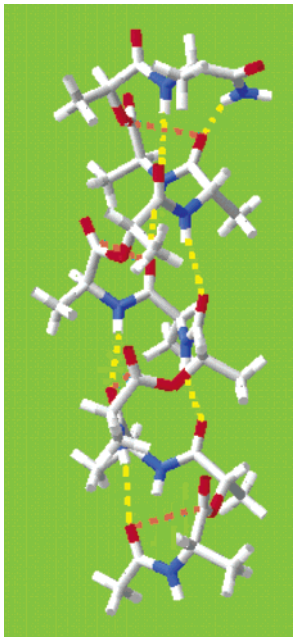


FIGURE 2. Charge polarization of the peptide bonds Δe (au) in 3_{10}-helices **I**, **II**, ***I4***, and ***II6*** (two depsipeptide mutants) at different levels of theory (unconstrained optimizations): (A) HF/3-21G values vs the B3LYP/6-31G* values (red circles); (B) HF/3-21G values vs the HF/6-31G** values (black squares).

CHART 2. AcALacA_2LacA_2LacA_2LacANH$_2$ *X2.5.8.11*, a Model for the α-Helix in 4-Ester CI2,[5] and the Selected Backbone Interactions[a]

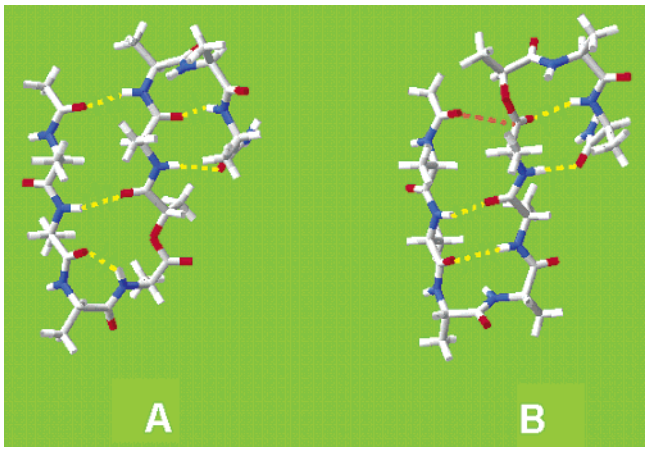


[a] The average C═O···HN bond distance in the array preceding the mutation sites (on the left-hand side of the diagram) is 1.997 Å compared to 2.107 Å in the parent helix **X**, in the array following the mutation sites it is 1.954 Å, compared to 1.972 Å in **X**. The average C_i═O···C_j═O (i = amide, j = ester) contact in ***X2.5.8.11*** is 2.765 Å (color-coded orange), compared to the average 2.910 Å in **X** (i = amide, j = amide), and 3.091 Å for the alternative ester contact in ***X2.5.8.11*** (i = ester, j = amide).

backbone-backbone H-bonds in the parent polyalanyl α-helix. The strengthening of these bonds could be due to to the increase in charge polarization of the peptide bonds as described earlier, section 3a.

c. Ala → Lac Substitutions in Hairpin Conformers. The results for the three fully optimized hairpin models are listed in Table 1 (**VI**, **VII**, **VIII**), and the examples of the structures are given in Chart 3. A wide

CHART 3. The Hairpin-based Triple-stranded *β*-sheet Models: the VII*5* (A) and VII*7* (B) Mutants, and the Selected Backbone Interactions[a]



[a] The numbering of residues (and mutation sites) begins at the N-terminus in the upper left corner of the diagram. The structure **VI** is obtained by removing three residues from the C-terminus, **VIII** by adding there one residue. The backbone−backbone C═O···HN bonds are shown in yellow. In ***VII7*** (B), the C_i═O···C_j═O bonding interaction is shown in orange.

range of the ΔE_{GT} values is also seen in this case. The substitution reactions are nearly thermoneutral in the case of the mid-turn linkages, i.e., peptide groups not involved in H-bonding at all (cf. entries **VI*4***, **VII*4***, **VIII*4***, **VII*8***, and **VIII*8*** in Table 1), which corroborates the results obtained for the single-strand α-helix turn **IX*2*,*3***.[31] However, several of the hairpin ΔE_{GT}'s are considerably larger than the helix ΔE_{GT}'s. The difference is quite conspicuous in the case of substitutions in several linkages that are exclusively N−H-acceptors, e.g., **VI*1***, **VII*1***, **VIII*1***, and **VIII*10***. Such magnitudes of the destabilization effects suggest an additional loss of bonding upon mutations in the *β*-sheetlike fragments, vide infra. On the other hand, the data also indicate a number of possible compensating backbone−backbone interactions in the depsipeptide hairpins which in several instances seem to be quite flexible. Thus, the substitution next to the midturn peptide linkage can induce a rotation of the backbone chain which brings the NH group not involved in any H-bond in the 'native' structure into the vicinity of the ester O group and the carbonyl O of the preceding peptide bond, see ***VII5*** in Chart 3A. This may have some compensating effect (cf. entries **VI*5***, **VII*5***, **VIII*5***, **VII*9***, and **VIII*9*** in Table 1). Yet another compensating interaction may result from the donor−acceptor contacts C_i═O···C_j═O as shown in Chart 3B (i = amide, j = ester; entries ***VII7*** and ***VIII7*** in Table 1): a rotation of the backbone chain enables the O···C_j═O approach in ***VII7*** at the distance of 2.950 Å, and the O···C_j═O angle of 100°, which is optimal for a bonding interaction.[32,33] The two effects seem to combine in the mutant ***VIII7***.

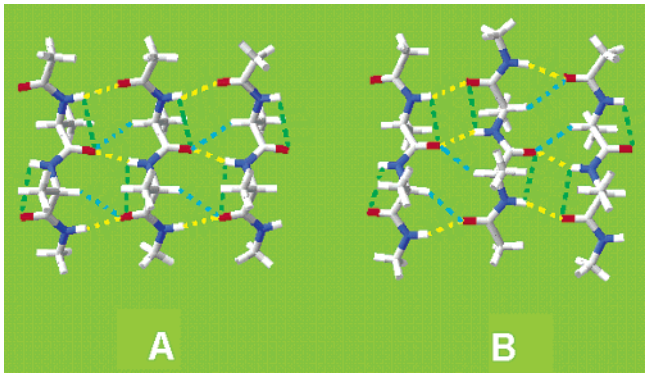
d. Ala → Lac Substitutions in Planar Parallel and Antiparallel *β*-Sheet Models. To avoid the complexity of the fully optimized *β*-hairpin models, we have examined several models assembled from the single strands which were kept in fixed conformations characteristic for the parallel and antiparallel *β*-sheets; see section 2. The results are summarized in Table 3. The layout of the table

TABLE 3. Group-Transfer Energies ΔE_{GT} (kcal mol^{-1}) for *m* (*m′*, *m″*) Ala → Lac Mutations of *N*-Acetyl Polyalanine Methylamides from the Isodesmic Reactions $AcA_nNHCH_3 + AcOMe \rightarrow AcA_xLacA_yNHCH_3 + AcNHMe$ (*m*, *m′*, and *m″* Denote the Mutation Sites) at the HF/3-21G[a] and B3LYP/6-31G*[b] Levels

		single strand	double-stranded β-sheet		triple-stranded β-sheet		
planar parallel	*m*	**XI**	**XII**		**XIII**		
tripeptide	**1**	3.3	11.0	4.9	14.4	13.7	5.3
AcA_2NHME	**2**	6.0	8.1	17.5	7.9	20.1	18.6
	3	3.2	10.1	4.3	12.6	13.1	4.7
planar parallel	*m*	**XIV**	**XV**				
tetrapeptide	**1**	3.4	12.3	5.3			
AcA_3NHME	**2**	5.9	7.3	16.7			
	3	5.8	16.6	8.1			
	4	3.3	4.7	11.2			
planar antiparallel	*m*	**XVI**	**XVII**		**XVIII**		
tripeptide	**1**	4.9	12.5	7.4	14.5	15.1	7.5
AcA_2NHMe	**2**	7.5	9.5	18.9	9.1	21.1	19.2
	3	2.2	12.7	3.3	13.9	14.8	3.6
planar antiparallel	*m*	**XIX**	**XX**	**XXI**			
tetrapeptide	**1**	5.5	7.6	12.9			
AcA_3NHMe	**2**	7.1	17.5	8.9			
	3	7.5	10.7	18.0			
	4	2.3	13.0	3.5			

[a] The total energies of the "wild-type" structures at the HF/3-21G level: **XI**, −734.5942875; **XII**, −1469.2446175; **XIII**, −2203.8988627; **XIV**, −979.0759374; **XV**, −1958.2272421; **XVI**, −734.5975173; **XVII**, −1469.2510859; **XVIII**, −2203.9087552; **XIX**, −979.0809868; **XX**, −1958.2430951; **XXI**, −1958.2311767 (hartrees). [b] The group transfer energies ΔE_{GT} (kcal mol^{-1}) at the B3LYP/6-31G* level for the selected structures: **XI*1-3***, 0.9, 3.3, 1.7; **XVI*1-3***, 1.4, 4.5, 2.2; **XVIII*2,3***, 5.8, 7.9; ***1′,2′***, 9.1, 13.9; ***1″-3″***, 4.9, 12.2, 2.2; **XIX*1-4***, 1.5, 4.5, 4.3, 2.5.

CHART 4. Planar Parallel and Antiparallel β-Sheet Models XIII (A) and XVIII (B) and the Selected Backbone Interactions[a]

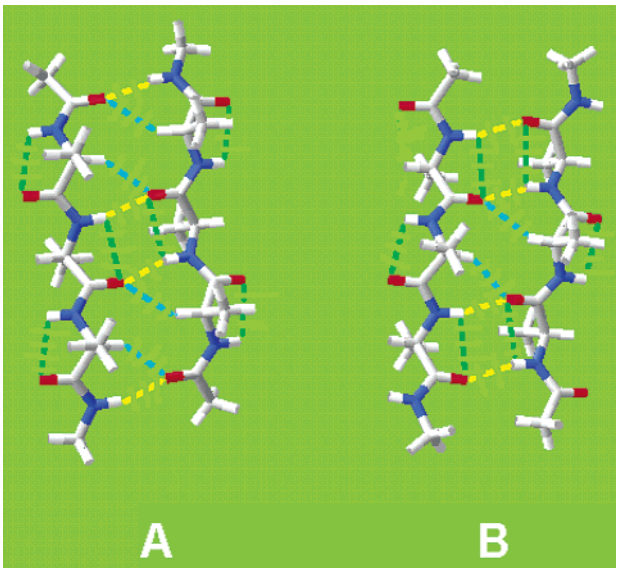


[a] The numbering of residues (and the mutation sites) begins in **XIII** at the H-bond donor N-terminus in the upper left corner of the diagram and continues at the top (N-terminus) of the neighboring (central) strand. The interstrand backbone−backbone C═O···HN bonds are color-coded yellow, the most important C═O···HC_α interactions are color-coded blue, and the most important intrastrand backbone−backbone C═O···HN bonds are shown in green.

is meant to reflect the topology of the model structures as shown in Charts 4 and 5.

For instance, the data for mutations of the triple-stranded parallel β-sheet model **XIII**, shown in Chart 4A, are listed beginning with the H-bond donor N-terminus (the upper-left corner of the chart) in the left-most column, the index *m* denoting the mutation site; the second column list the data for the central strand, beginning at the top again, with the index *m′* etc. Consequently, the entries in a row represent values for a perpendicular array of H-bonded peptide bonds, labeled *m*, *m′*, *m″*. Only one column of the data is given for the

CHART 5. Planar Antiparallel β-Sheet Models XX (A, the LSL Sequence of the H-Bonded Rings) and XXI (B, the SLS Sequence), and Selected Backbone Interactions[a]



[a] In both cases, the numbering begins at the N-terminus in the upper left corner; the two strands are related by the 2-fold axis. The interstrand backbone−backbone C═O···HN bonds are color-coded yellow, the most important interstrand C═O···HC_α interactions are color-coded blue, and the most important intrastrand backbone−backbone C═O···HN bonds are shown in green.

structures **XX** and **XXI** shown in Chart 5A and Chart 5B, respectively, since the two strands in these alternative antiparallel tetrapeptide complexes are related by the 2-fold axis (note the sequence LSL in Chart 5A (**XX**) and SLS in Chart 5B (**XXI**), S and L referring to the small and large H-bonded rings).

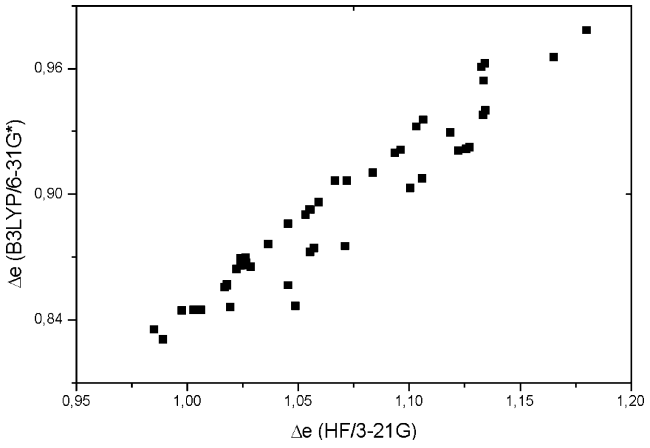


FIGURE 3. Charge polarization of the peptide bonds Δ*e* (au) in the antiperiplanar *β*-sheet model **XVIII**, its two depsipeptide mutants **XVIII*2*** and **XVIII*2′***, and in strands **XIX** (Table 3) and **IX**[31] (constrained optimizations) at the HF/3-21G and B3LYP/6-31G* levels of the theory.

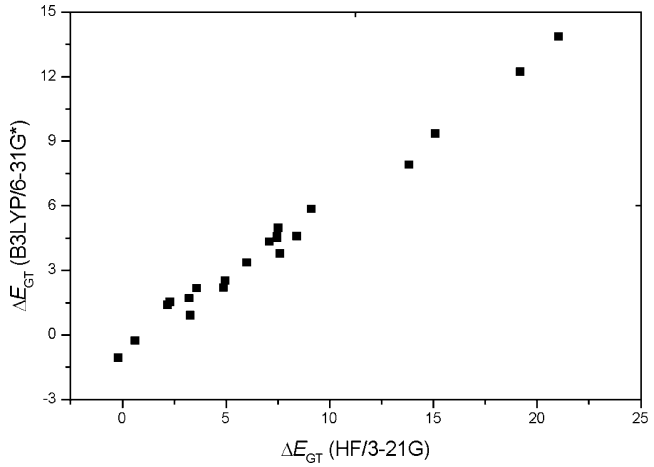


FIGURE 4. Group-transfer energies ΔE_{GT} (kcal mol^{-1}) for the 3_{10}-helices **I*4***, **II*6***, the antiperiplanar *β*-sheet model mutants **XVIII*2*,*3*,*1′*,*2′*,*1″*-*3″***, and single-strand mutants **XI*1*-*3***, **XVI*1*-*3***, **XIX*1*-*4*** (cf. Table 3, and footnote b), and **IX*2*,*3***[31] at the HF/3-21G and B3LYP/6-31G* levels of the theory.

The range of the ΔE_{GT} values is significantly greater for the present *β*-sheet models than for the 3_{10}- and α-helix models. There could be two major reasons for that increase. In contrast to the results obtained for the non-H-bonded peptide linkages in the α-helix turn or *β*-hairpin, Ala → Lac mutations in the single *extended* strands **XI**, **XIV**, **XVI**, and **XIX**, are considerably endothermic (Table 3), hence the large destabilization effects in the *β*-sheet fragments are due in part to loss and weakening of the *intrastrand C═O···H−N bonding*.[36] The second destabilizing contribution could be due to loss of the *interstrand C═O···H−C*$^{\alpha}$ *bonding*. Numerous structural data indicate that there is an additional interstrand bonding in the *β*-sheets between the peptide carbonyl O and the C$^{\alpha}$−H;[37,38] the strength of such interactions was estimated to approach 1.1−2.6 kcal mol^{-1} in formamide and *N*-methylacetamide dimers.[39,40] The most important interactions of the two types are indicated in Charts 4 and 5 in addition to the regular C═O···H−N bonds between the adjacent peptide chains. The corresponding distances, e.g. in the fully optimized structure **VIII**, are 2.370−2.448 Å for the C═O···H−C$^{\alpha}$ interactions, and 2.186−2.382 Å for the "intrastrand" C═O···H−N interactions (HF/3-21G).

The charge-polarization parameters Δ*e* for the peptide bonds in several 3_{10}-helix structures obtained at the HF/3-21G and B3LYP/6-31G* levels were compared in Figure 2. The analogous plot for the present *β*-sheet models shows a somewhat greater scatter, Figure 3, but the correlation supports the earlier conclusion that the HF/3-21G model is qualitatively useful in reproducing charge distribution in the peptide chains.

4. Comparison of the Group-Transfer Energies at the HF/3-21G and B3LYP/6-31G* Levels

Since the HF level of theory using the 3-21G basis set significantly overestimates the energy of hydrogen bonds,[41] the question arises how reliable are the values listed in Tables 1 and 3. To address this question, reoptimizations at the B3LYP/6-31G* level were carried out for a number of structures which define the obtained range of the group-transfer energies—the two 3_{10}-helices (**I**, **II**), the triple-stranded antiperiplanar *β*-sheet (**XVIII**), and single strands in the extended (**XI**, **XVI**, **XIX**) and helical (**IX**) conformations,[31] along with the corresponding depsipeptide mutants. All these data are combined in the plot in Figure 4. The correlation with the slope of 0.67 is quite satisfactory. The slope reflects the changes in the ΔE_{GT} values in the extended strands (Table 3, footnote b), whereas the actual scaling back of the HF/3-21G ΔE_{GT}'s in several cases of the 3_{10}-helix mutations is closer to 0.5.

5. Origin of the Destabilization Effects; Partitioning of the Group-Transfer Energy into Lost H-Bonds, Decrease in Lewis Basicity of the C═O Group, and O···O Repulsion

It seems reasonable to assume that the variation in charge polarization of the peptide bonds and in the CO··· HN separation reflect the variation in the strength of backbone H-bonding.[42,43] It is therefore significant that the destabilizing effects of Ala → Lac mutations, i.e., the group-transfer energies, correlate with the degree of charge polarization of the mutated peptide linkages measured by the difference Δ*e* in Mulliken populations at H and O in HN−C═O, and with with the H-bond

(36) Shamovsky, I. L.; Ross, G. M.; Riopelle, R. *J. Phys. Chem. B* **2000**, *104*, 11296.
(37) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83. Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248. Bella, J.; Berman, H. M. *J. Mol. Biol.* **1996**, *264*, 734.
(38) Lee, K. M.; Chang, H.-C.; Jiang, J.-C.; Chen, J. C. C.; Kao, H.-E.; Lin, S. H.; Lin, I. J. B. *J. Am. Chem. Soc.* **2003**, *125*, 12358.
(39) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750.
(40) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963.
(41) Frisch, M. J.; Del Bene, J. E.; Binkley, J. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 2279. Cramer, C. J. *Essentials of Computational Chemistry*; John Wiley & Sons: Chichester, 2002; pp 179−181.
(42) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997. Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48.
(43) Harris, T. K.; Mildvan, A. S. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 275. Kang, Y. K. *J. Chem. Phys. B* **2000**, *104*, 8321.

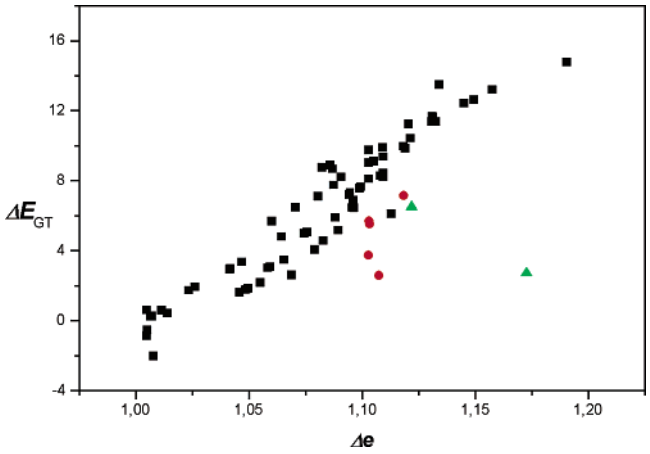


FIGURE 5. Dependence of the group-transfer energy ΔE_{GT} (kcal mol^{-1}) for *m*A → Lac mutations of *N*-acetyl polyalanyl amides on the difference Δe (au) in H and O Mulliken populations of the substituted *m* peptide bond HNC═O: the data from the unconstrained optimizations of the helical and hairpin conformers (Table 1). The color-coded data sets represent mutants with compensatory backbone interactions resulting from the *i*, *i* + 3 ↔ *i*, *i* + 4 transitions (red) or donor−acceptor contacts/H-bonding (green).

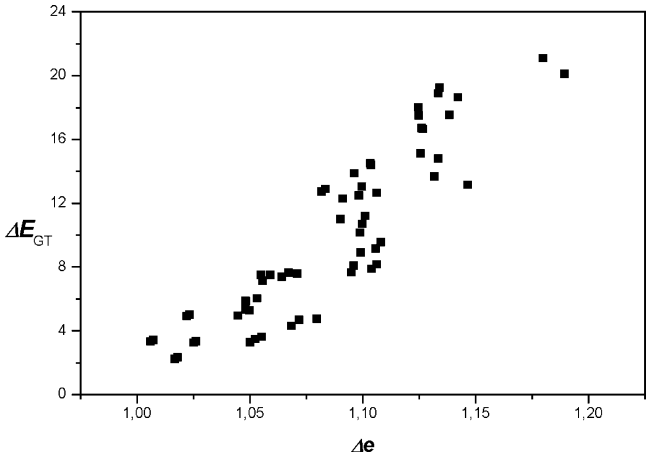


FIGURE 6. Dependence of the group-transfer energy ΔE_{GT} (kcal mol^{-1}) for *m*A → Lac mutations of *N*-acetyl polyalanyl methylamides on the difference Δe (au) in H and O Mulliken populations of the substituted *m* peptide bond HN−C═O: the data from constrained optimizations (see Computational Methods) of the planar parallel and antiparallel β-sheet models (Table 3).

distances in the "wild-type" structures. The ΔE_{GT} vs Δe plot in Figure 5 shows a reasonable correlation for the helix and hairpin mutants with the preserved "native" structure and pattern of backbone-backbone interactions. For the remaining mutants (the data marked with asterisks in Table 1, and color-coded in Figure 5), the ΔE_{GT} values deviate from the overall distribution in a manner consistent with attenuation of the loss of bonding as described earlier.

A similar plot is obtained for the constrained β-sheet models, Figure 6.

Finally, the dependence of the group-transfer energies on the H-bond distances can be tested on the samples including the peptide bonds involved exclusively either as N−H-donors or N−H-acceptors. The plots for the

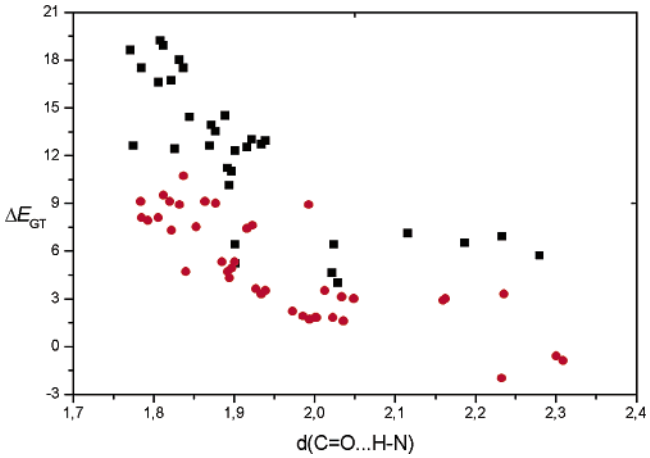


FIGURE 7. Dependence of the group-transfer energy ΔE_{GT} (kcal mol^{-1}) on the H-bond distance in the "wild-type" structure (Å): (A) black squares represent substitutions in the peptide linkages that are exclusively NH-bond donors; (B) red circles represent substitutions in the peptide linkages that are exclusively NH-bond acceptors. The ΔE_{GT} data are taken from both Tables 1 and 3.

combined data from Tables 1 and 3 are shown in Figure 7: in each ΔE_{GT} sample, the correlation does seem to account for a major fraction of variance.

The correlations shown in Figures 5−7 suggest that loss or weakening of a hydrogen bond do usually constitute a major contribution to the destabilization effect of the Ala → Lac mutation, and the partitioning of the group-transfer energies should yield reasonable estimates of the electronic association energy of the backbone-backbone H-bonds. To discuss such a partitioning, it is convenient to distinguish four categories of the peptide linkages that are mutated in this study: (i) N−H-acceptor bonds; (ii) N−H-donor bonds; (iii) bonds that are simultaneously N−H-acceptors and N−H-donors; (iv) bonds that do not participate in H-bonding.

In the first case, the mutations do not remove any C═O···H−N bonds, but merely weaken the extant ones. The destabilization is expected because of the difference in the dipole moments and Lewis basicity of the esters and the amides,[44] but the effect cannot be very large. Abraham's H-bond structural group constants[45] predict ∼1−2 log unit difference in 1:1 complexation constants in the gas phase.[46] This estimate is quite consistent with the magnitude of the quoted above ΔE_{GT}'s for **I*1***−**IV*1***, considering that the HF/3-21G values of the group-transfer energies scale back at the higher level of theory, vide supra, by the factor of 0.5−0.7. As was already mentioned, the much larger destabilization effects resulting from substitutions in the β-sheet N−H-acceptor peptide linkages suggest an additional loss of bonding. One of such additional contributions could be the intra-strand H-bonding which is quite large in the extended single strands, see Table 3: the corresponding distances

(44) Ethyl acetate DN = 17.1, *N*-methyl acetamide DN = 26.6: Gutmann, V. *The Donor−Acceptor Approach to Molecular Interactions*; Plenum Press: New York, 1978. Ethyl acetate μ = 1.90 D, *N*-methyl acetamide μ = 4.42 D: Vogel, P. *Chimie Organique*; De Boeck Université: Bruxelles, 1997.

(45) Abraham, M. H.; Platts, J. A. *J. Org. Chem.* **2001**, *66*, 3484.

(46) Marco, J.; Orza, J. M.; Notario, R.; Abboud, J.-L. M. *J. Am. Chem. Soc.* **1994**, *116*, 8841.

in the fully optimized structure **VIII** are 2.186−2.382 Å for the "intrastrand" C═O···H−N interactions. In the case of the constrained β-sheet models, this contribution can be separated since by definition (Hess's Law), the loss of electronic association energy ΔD_e due to an Ala → Lac mutation is equal to the difference of the group-transfer energies for the binary or ternary complex and the corresponding single extended strand $\Delta D_e = \Delta E_{GT}(SSm) - \Delta E_{GT}(Sm)$. Some of the energies obtained this way are slightly larger than the ΔE_{GT}'s for the helix mutants **I*2*-V*2***, which is perhaps an indication of a small contribution of the interstrand C═O···H−C^α interactions, see section 3d. Thus, the overall range of the destabilization effects due to the decrease in C═O basicity is 1.1−2.6 kcal mol^{-1} (HF/3-21G; B3LYP/6-31G* estimate 0.7−1.7 kcal mol^{-1}).

In the second case, the Ala → Lac mutations in the helical N−H-donor bonds remove one H-bond and introduce a close O···O contact, while the mutations in the β-sheet N−H-donor bonds in addition remove an intrastrand H-bonds. Thus, the ΔE_{GT}'s for mutants **I*6*,*7***, **II*8*,*9***, **III*8*,*9*,*10***, **IV*12*,*13*,*14***, and **V*9*,*10*** (Table 1) and the $\Delta D_e = \Delta E_{GT}(SSm) - \Delta E_{GT}(Sm)$ for the mutants **XII*2′***, **XIII*2″***, ***XV3*,*2′***, **XVII*2′***, **XVIII*2″***, **XX*2***, and **XXI*3*** (Table 3) primarily comprise the loss of H-bond and the O···O repulsion. By analogy to the experimental approach,[19] an estimate of the latter contribution can be based on a comparison of the group-transfer energies for Ala → Lac mutations (N−H → O replacement) with the group-transfer energies for Ala → Iba (isobutyric acid) mutations which substitute the peptide linkage with the keto methylene moiety: $AcA_nNHR + MeCOCH_2Me \rightarrow AcA_x$-$CH_2CH(CH_3)COA_yNHR$ + MeCONHMe (R = H, CH_3) (N−H → CH_2 replacement, the group-transfer reaction between a peptide and ethylmethyl ketone). The corresponding optimizations (HF/3-21G) were made for the Ala → Iba mutations **I*5C*** (11.1), **I*7C*** (7.8), **III*7C*** (11.6), **III*9C*** (10.1), **IX*2C*** (6.6), **IX*3C*** (7.2), **XVI*2C*** (5.6), **XIX*2C*** (5.4), **XIX*3C*** (5.6), and ***XVII2′C*** (13.2) (all bracketed values in kcal mol^{-1}). The substitutions in the single helix turn **X** are in this case highly endothermic, av. 6.9 kcal mol^{-1}, compared to 0.2 kcal mol^{-1} for the Ala → Lac mutations;[31] the substitutions in the single extended strands **XVI** and **XIX** are slightly less endothermic than the Ala → Lac mutations, av. 5.5 kcal mol^{-1}. If these values are used to "correct" the group-transfer energies quoted above in the brackets, and the resulting figures are subtracted from the Ala → Lac ΔE_{GT}'s, the differences average to 3.0 kcal mol^{-1} in the case of the 3_{10}-helices, where the O···O separation is 3.5−3.7 Å, see section 3a, and to 3.6 kcal mol^{-1} in the case of the α-helices and β-sheets, where the O···O separation is 3.0−3.2 Å. Thus, the B3LYP/6-31G* estimate of the O···O repulsion effect, see section 4, is 1.5−2.4 kcal mol^{-1}. Boger et al. estimate of the $\Delta\Delta G°$ due to O···O repulsion is 2.6 kcal mol^{-1} from a comparison of the vancomycin and vancomycin aglycon binding affinity for $Ac(AcK)^DA^DA$ and $Ac(AcK)^DA^DLac$ vs $Ac(AcK)^DAIba$, the latter ligand incorporating CH_2 in place of the amide NH.[19]

In the case of the third category, the Ala → Lac mutations at the peptide bonds that are simultaneously N−H-acceptors and N−H-donors, the destabilization effect comprises three major contributions in a helix, and four major contributions in a β-sheet. With the above estimates of these contributions in hand, and assuming that the effect of C═O basicity is slightly larger for the peptide bonds in the helix interior than in its first turn (i.e. 2.0 kcal mol^{-1} in the 3_{10}-helix, 2.5 kcal mol^{-1} in the α-helix, HF/3-21G), one can now estimate the electronic binding energies of the backbone-backbone C═O···HN interactions. Using the data for the mutations **II*6***, **IV*9***, and **XIII *2″***, and scaling the results to the B3LYP/6-31G* level, the following *maximum* values are obtained for the presently examined set of the secondary structure models: 3_{10}-helix $D_e = -1.7$ kcal mol^{-1}, α-helix $D_e = -3.8$ kcal mol^{-1}, and β-sheet $D_e = -6.1$ kcal mol^{-1}.

The observed trend in estimated D_e's is consistent with the earlier results of quantum-mechanical calculations on the isolated amide model systems[47] and with the available spectroscopic and structural evidence.[22] In contrast, molecular mechanics calculations yield reverse ordering of the α-helix and β-sheet H-bonding energies.[48] This might be a consequence of using the same partial charges for both folds; it has also been suggested that the force fields using atom-centered partial charges cannot give reliable dependence of H-bonding energetics on geometry.[49]

It should be noted that D_e's based on the group-transfer energies do not seem to capture the entire stabilization effect of H-bond formation in the complexes of polypeptides. For instance, the B3LYP/6-31G* data for **XVI*2*** and **XVIII*2″*** give β-sheet $D_e = -7.7$ kcal mol^{-1} without any correction for O···O repulsion, and $D_e = -5.3$ kcal mol^{-1} with such a correction. On the other hand, the average values $D_e = -7.8$ or -8.6 kcal mol^{-1} are obtained using the total energies of **XVIII** and **XVI** or the total energies of **XVIII** and its isolated strands (from single point calculations), respectively. Assuming that the repulsion effect is properly accounted for, the difference can perhaps be attributed to the interstrand C═O···H−C^α interactions, see Section 3d, and to the stabilization of the individual strands upon formation of the three-stranded β-sheet. The latter possibility highlights the difficulty in isolating the "hydrogen bonding" energy in the case of interaction of two polar groups embedded in a polypeptide chain and will be discussed elsewhere.

6. Conclusions

Site-directed mutagenesis provides a wealth of experimental information which is often inherently difficult to interpret. This is particularly true in regard to the problem of secondary structure stability, since the local interactions might include nonclassical contributions (secondary bonding) which cannot be examined without quantum-mechanical calculations. One possible way to approach this problem is to describe a mutation in terms of an isodesmic equation. It seems now, at least in the case of the Ala → Lac substitutions, that this approach can indeed be successful. Heats of the group-transfer reactions between peptides and methyl acetate, ΔE_{GT} (Scheme 1), are found to be quite sensitive to the features

(47) Mitchell, J. B. O.; Price, S. L. *J. Comput. Chem.* **1990**, *11*, 1217.
(48) Lazaridis, T.; Archontis, G.; Karplus, M. *Adv. Protein Chem.* **1995**, *47*, 231. Sheu, S.-Y.; Yang, D.-Y.; Selzle, H. L.; Schlag, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12683.
(49) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908.

of secondary structure. Destabilizing effects of such mutations are absent in type I turns, small in the 3_{10}-helices, somewhat greater in the α-helices, and quite large in the *β*-sheet models; ΔE_{GT}'s also tend to increase upon extension of the H-bonded array of peptide linkages.[50] Qualitatively, this picture is independent of the level of the theory employed since the HF/3-21G and B3LYP/6-31G* ΔE_{GT}'s show a good correlation with the slope of 0.67, cf. Figure 7. In several depsipeptide mutants, the "wild-type" H-bonding pattern and molecular geometry are distorted by compensating backbone−backbone interactions. In all other cases, the group-transfer energies correlate with the H-bond distances in the "wild-type" structures, and with the charge-polarization indices of the mutated peptide linkages. Partitioning of the group-transfer energies into loss of interstrand and intrastrand H-bonds, decrease in C═O basicity, and O···O repulsion, yields the following B3LYP/6-31G* estimates of the maximum electronic association energies of the backbone-backbone C═O···HN bonds in the presently examined model structures: 3_{10}-helix $D_e = -1.7$ kcal mol^{-1}, α-helix $D_e = -3.8$ kcal mol^{-1}, and *β*-sheet $D_e = -6.1$ kcal mol^{-1}. These assessments will be affected by inclusion of the medium and specific solvation effects, the vibrational contributions, and the improvement of the theoretical model. Nonetheless, such corrections seem unlikely to alter in any fundamental way the emerging picture of considerable variation in charge polarization of the peptide bonds, and in the stabilization that the backbone interactions provide, in different elements of secondary structure. Thus, our findings corroborate the recent proposition that electronic configuration of the peptide bonds can vary along the amide rehybridization/polarization path, and that the optimal stability of a given element of secondary structure requires a specific location of its peptide bonds along this path;[51] since electronic configuration of a peptide bond depends inter alia on the inductive and hyperconjugative side chain−peptide bond interactions, secondary structure stability may also depend on electronic properties of the amino acid side chains.

Acknowledgment. We thank the administration of the High-Performance Computer Center at Bilkent University for the generous access to the center facilities. The parallelized version of Gaussian 98, Revision A.7,[30] was installed on the Sun Enterprise 4500 Server by Professor Ulrike Salzner of the Department of Chemistry, Bilkent University.

Supporting Information Available: Gaussian98 archive entries (the Cartesian coordinates and the total energies) for the peptide, depsipeptide, and auxiliary structures. This material is available free of charge via the Internet at http://pubs.acs.org.

JO0358372

(50) This is consistent with the view which attributes cooperativity in the helix formation to the strengthening of the backbone-backbone H-bonds (see, for instance: Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 945; and ref 28b); this view has recently been questioned, ref 29a. For the evidence that *β*-sheet formation is cooperative in perpendicular direction see: Schenck, H. L.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4869. For the opposing view: Zhao, Y.-L.; Wu, Y.-D. *J. Am. Chem. Soc.* **2002**, *124*, 1570.

(51) Cieplak, A. S. In *The Amide Linkage: Selected Structural Aspects in Chemistry, Biochemistry and Materials Science*; Greenberg, A., Breneman, C., Liebman, J. F., Eds.; John Wiley & Sons: New York, 2000; Chapter 17, pp 565−597. Cieplak, A. S. *Chem. Rev.* **1999**, *99*, 1265. Cieplak, A. S. *Struct. Chem.* **1994**, *5*, 85. Cieplak, A. S. *J. Am. Chem. Soc.* **1985**, *107*, 271.